



Part 2: Question & Answer Session

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Lola Fatoyinbo (lola.fatoyinbo@nasa.gov) or Abigail Barenblitt (abigail.barenblitt@nasa.gov).

Question 1: How do you deal with the slc error of Landsat 7 between 2003 and 2014?

Answer 1: GEE masks the SLC error for Landsat 7, so for a single image you will see stripes of missing imagery. To compensate for this we take a mosaic of several Landsat 7 images.

Question 2: How can we unfold the "black box" and prepare our own algorithm according to our own need?

Answer 2: Understanding the black box and preparing new machine learning algorithms requires a deeper understanding of machine learning. Our work is more interested in using models that we already know are robust.

Question 3: Is there a reason you choose QGIS 3.10 vs 3.14?

Answer 3: You can use either. We chose 3.10 because that is what I was working with and we wanted to make sure there was consistency. However, a participant emailed to say they were able to install the plugin in 3.14.

Question 4: I've some doubts in exporting the mangrove's data from GEE. I'm trying to export the data with the spatial resolution of 10 m. The GEE shows an error that it cannot be exported because of greater pixel size. So what should I do to reduce the pixel size?

Answer 4: If you are working with Landsat, the spatial resolution is 30m, so I would export at that resolution to match. If you are working with data at a smaller resolution, you may need to export your data in batches because of GEE capacity.

Question 5: Can we use this method to predict future extent/distribution?



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Answer 5: You could potentially extrapolate trends of mangrove loss from this analysis, but it would not be super robust. This method relies on available optical imagery rather than modeled projections like climate models. The drivers of loss vary.

Question 6: Hello As you write that code in GEE website as JavaScript but in QGIS with GEE plugin, it will be python so in this case you could provide the GEE code in python too?

Answer 6: I do not have this code in python since I primarily export from GEE to QGIS and work with the user interface. There are capabilities for using GEE directly through QGIS as you mention, but it is not something I have been able to explore yet.

Question 7: What is the median reducer? Is it a filter? Can any other filter(s) be used to remove cloud contamination?

Answer 7: The median reducer will grab the median values for that pixel. That helps create a cleaner, cloud-free mosaic. You can also use `.qualityMosaic()` to get the default “best” pixel value.

Question 8: Can you please explain again why FCC (False Color Composite) is more useful than TCC (True Color Composite) for identifying mangroves?

Answer 8: You can really use either, but since we are distinguishing between mangroves and other areas of vegetation, FCC makes it easier to visualize areas of dense vegetation in red, which in turn makes it easier for us to make training data.

Question 9: The instructor is adding and removing comments (`//`) from lines of code to run portions of the code. How is this done?

Answer 9: If you hit Command + `'` on a Mac or Ctrl + `'` on a PC

Question 10: In Sundarbans Mangrove of Ganges-Brahmaputra delta region elevation showing in SRTM 0 or negative Zvalue, so how to correct it?

Answer 10: I quickly checked in GEE and was not having this issue. I would have to look at your code to say more.

Question 11: Is it better to use points or polygons for the classification?



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Answer 11: Polygons are great if there are areas that are uniform and make it easier to collect more training data quickly. Points can be helpful for refining the model. But really you can use either.

Question 12: In general, how many pixels should a training area cover?

Answer 12: It's difficult to say. Generally you start with a few, test the model, and add more as you refine the model. You don't want to start with too much training data at first. Start small and refine as you go.

Question 13: Can you explain the Random Forest model again? Is it sort of a hierarchical clustering?

Answer 13: The model works through a series of decision trees (essentially a hierarchy). There are several trees for each pixel and each tree "votes" on the correct classification for a pixel. The majority of votes decide what the pixel is classified as. If the majority of trees conclude Mangroves, then that is what will be used in the model.

Question 14: How do you know if you have too much training data?

Answer 14: GEE will time out on you or give you an error for Memory limit exceeded.

Question 15: How do we know the dates of the satellite image we're using? That is to say that day and month of 2009 or 2019? Or is it an annual average?

Answer 15: The method we use pulls every available image from the study period we specify. When we use `.median()` and `.qualityMosaic()` we are essentially filtering through all available images and pulling the best pixel from that time period on a pixel by pixel basis across the study area.

Question 16: I guess all the signatures were made once for both pre-post years, and then the same training dataset was used for classification in 2009 and 2019. Does this google earth engine base-map (satellite), switches back to 2009 when I call my image of 2009. I guess, is it okay to re-draw signatures for 2009 separately for the purpose of classification too?

Answer 16: I'm not sure I understand this question. If you are referring to the training data, we create training data specific to each year of interest. When it comes to the base-map, we don't want to rely on this for the classification since we don't know what year it's from.



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Question 17: Could we select a location (i.e. using polygons and points) and then the computer - engine tells us the possible bands that we can select? Instead of telling the model the bands a priori?

Answer 17: I don't know a way to do this ahead of time. However, after running the model you can use `.explain()` to get a print out of how much each band contribute to the model and decide if you need to change the inputs from there.

Question 18: How is python used instead of Javascript in Earth Engine?

Answer 18: There is a python plug-in for GEE but it is still a bit clunky. You can use the QGIS GEE plug-in to work in python, but I find that Javascript is the best option for now.

Question 19: Why did you create a L7 mask instead of a L8 mask?

Answer 19: For the second part of the tutorial we were looking at 2009 imagery, which is not available through L8. Therefore, we have to use L7.

Question 20: When we cloud mask the collection, how likely is it the median composite would have the masked pixel?? If still there exists masked pixels, how do we resolve it?

Answer 20: You can still have masked pixels. One way to work around this is to expand the study period.

Question 21: As the median composite may have different bands with different timestamps, would that impact classification? If yes, how to address it?

Answer 21: Because we are pooling the median, typically we assume it will be close enough to our time period. Try to work within a smaller time period and expand out if you need to.

Question 22: Can we use GEE to determine the future pathways of a certain scenario (example: BAU vs. Green Economy scenario)?

Answer 22: Yes, GEE allows you to import future scenario model outputs.

Question 23: Is it possible to use pan-sharpened imagery to generate a classification map? What would be the advantages and disadvantages?



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Answer 23: Yes. Generate the training data from the high resolution imagery. But these tend to be larger files and there may not be an ample amount of data.

Question 24: Is it possible to perform trend analysis of Mangrove loss/gain in GEE?

Answer 24: Yes, we will be doing this in Part 3 next week. In the code from today, you can get an extent map from each year and compare.

Question 25: How did you uncomment several lines of code?

Answer 1: You can highlight several lines and then click Command + '/' on a Mac or Ctrl + '/' on a PC

Question 26: Is there a plugin to do all of this inside QGIS?

Answer 26: You have to download the plugin. You can find the link on the ARSET website

Question 27: My question is what does google earth engine run on mobile phone or only on laptop?

Answer 27: You can run GEE on a phone but it's not user friendly

Question 28: How to plot also the 2019 composite with "real color" values?

Answer 28: You would select the Red, Green, and Blue bands (B4, B3, B2)

Question 29: Do you use AWEI (Automated Water Extraction Index) to analyze mangroves or changes in the Earth's surface?

Answer 29: No, but you could easily add this to the list of predictors in the random forest assessment

Question 30: How do we get the image from GMW?

Answer 30: At the top of the code editor, you can use the search bar to search for Global Mangrove Watch

Question 31: If mangroves are encompassed within the definition of forest in a host country, should the definition of forest be used as a minimum mapping unit (MMU) and, if so, should the MMU be applied to the classification output rather than to the original satellite image?



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Answer 31: We try to separate out mangrove from forest by focusing on coastlines at low elevation. However, you could use a defined MMU to clean up the final model. This is a little harder with mangroves because you are trying to separate out two types of vegetation, but you may decide to use high res or other background imagery to clean up the final model. The random forest model does not have to be a final product and we usually do some sort of post-processing to get a final map that we feel is accurate.

Question 32: Is the NDVI for the whole year or for a particular date? Does it give a 1 year result?

Answer 32: We pick the NDVI value for the pixel that was chosen as the “best” pixel from our landsat composite.

Question 33: About the selection of 80/20% for training/testing, is that selection within the polygons already chosen as mangrove/no-mangrove?

Answer 33: Yes, this is from within the training data we generated.

Question 34: How can I map the 2009 and 2019 Landsat composites so I can compare them with the classified Landsat images?

Answer 34: In the code you will find lines for `Map.addLayer()`, which add the Landsat composite as a layer to the map

Question 35: In the 'Begin Random Forest Classification' step, what's your reasoning for selecting 100 trees and 5 predictors? What are the pros and cons of higher vs. lower values for these?

Answer 35: You can play around with these numbers and this is one of the ways we can work on improving our map. In previous models I've run for mangrove extent mapping these were the numbers that worked for me.

Question 36: What are the implications of creating a median value composite over a multi-year window? Does this only work here because seasonality does not affect mangrove vegetation? e.g. for a deciduous forest would you want to only composite certain seasons?

Answer 36: That is correct. The median composite works well in tropical areas where we experience less seasonality. In deciduous forests you would want to create landsat composites from only the growing season.



Remote Sensing for Mangroves in Support of the UN Sustainable Development Goals

November 5 - 19, 2020

Question 37: When doing your accuracy assessment is it important to be looking at imagery you know is from the same time roughly? I assume you would not want to be doing your assessment of the accuracy of the 2009 classification results with a generic image service that may be from current imagery.

Answer 37: Yes. You should try to acquire imagery from as close to the year of the classification as you can get. Otherwise you won't have a true assessment of the accuracy. I recommend using high res data of Google Earth Pro imagery.

Question 38: Is there big differences in performance between running the scripts in Java and Python? Why did you choose Java?

Answer 38: GEE runs primarily in JavaScript. There are Python plugins available in GEE, but they are still improving. We use JavaScript because that's what this particular platform requires. There shouldn't be much of a difference in performance if you ran a random forest in Python, but you would be using a different platform most likely.